

# Risk assessment of cancer patients based on HLA-I alleles, neobinders and expression of cytokines

Anjali Dhall, Sumeet Patiyal, Harpreet Kaur, and Gajendra P. S. Raghava\*

Department of Computational Biology, Indraprastha Institute of Information Technology,  
Okhla Phase 3, New Delhi-110020, India.

## Mailing Address of Authors

Anjali Dhall (AD): <a href="mailto:anjaliid@iiitd.ac.in">anjaliid@iiitd.ac.in</a>	ORCID ID: <a href="https://orcid.org/0000-0002-0400-2084">https://orcid.org/0000-0002-0400-2084</a>
Sumeet Patiyal (SP): <a href="mailto:sumeetp@iiitd.ac.in">sumeetp@iiitd.ac.in</a>	ORCID ID: <a href="https://orcid.org/0000-0003-1358-292X">https://orcid.org/0000-0003-1358-292X</a>
Harpreet Kaur (HK): <a href="mailto:hks04180@gmail.com">hks04180@gmail.com</a>	ORCID ID: <a href="https://orcid.org/0000-0003-0421-8341">https://orcid.org/0000-0003-0421-8341</a>
Gajendra P. S. Raghava (GPSR): <a href="mailto:raghava@iiitd.ac.in">raghava@iiitd.ac.in</a>	ORCID ID: <a href="https://orcid.org/0000-0002-8902-2876">https://orcid.org/0000-0002-8902-2876</a>

## \*Corresponding Author

Prof. Gajendra P. S. Raghava  
Head and Professor  
Department of Computational Biology  
Indraprastha Institute of Information Technology, Delhi  
Okhla Industrial Estate, Phase III  
New Delhi, India – 110020 Office: A-302 (R&D Block)  
Phone: 011-26907444  
Email: [raghava@iiitd.ac.in](mailto:raghava@iiitd.ac.in)  
Website: <http://webs.iiitd.edu.in/raghava/>

## Abstract

Advancements in cancer immunotherapy have shown significant outcomes in treating various types of cancers. In order to design effective immunotherapy, it is important to understand immune response of a patient based on its genomic profile. We compute prognostic biomarkers from 8346 cancer patients for twenty types of cancer. These prognostic biomarkers has been computed based on i) presence of 352 human leucocyte antigen class-I (HLA-I), ii) 660959 tumor-specific HLA1 neobinders and iii) expression profile of 153 cytokines. It was observed that survival risk of cancer patients depends on presence of certain type of HLA-I alleles; for example LIHC cancer patients with HLA-A\*03:01 are on lower risk. Our analysis indicate that neobinders of HLA-I alleles have high correlation with overall survival of certain type of cancer patients. For example HLA-B\*07:02 binders have 0.49 correlation with survival of LUSC and -0.77 with KICH cancer patients. It is clear from above analysis that HLA and their binders have major role in survival of cancer patients suffering from different types of cancer. In addition, we compute prognostic biomarkers for 20 types of cancer based on each type of cytokine expression. Higher expression of few cytokines is survival favourable like IL-2 for BLCA cancer patients whereas IL-5R survival unfavourable for KICH cancer patients. In order to facilitate scientific community we developed a web-based platform CancerHLA1 that maintain raw and analyzed data (<https://webs.iitd.edu.in/raghava/cancerhla1/>).

**Keywords:** HLA-alleles, neobinders, cancer, prognostic biomarkers, cytokines

## List of Abbreviations

- **BLCA**- bladder urothelial carcinoma
- **BRCA**- breast invasive carcinoma
- **CESC**- cervical squamous cell carcinoma and endocervical adenocarcinoma
- **CHOL**- cholangiocarcinoma
- **GBM**- glioblastoma multiforme
- **HNSC**- head and neck squamous cell carcinoma
- **KICH**- kidney chromophobe
- **KIRC**- kidney renal clear cell carcinoma
- **KIRP**- kidney renal papillary cell carcinoma
- **LIHC**- liver hepatocellular carcinoma

- **LUAD**- lung adenocarcinoma
- **LUSC**- lung squamous cell carcinoma
- **OV**- ovarian serous cystadenocarcinoma
- **PAAD**- pancreatic adenocarcinoma
- **PRAD**- prostate adenocarcinoma
- **READ**- rectum adenocarcinoma
- **SKCM**- skin cutaneous melanoma
- **STAD**- stomach adenocarcinoma
- **THCA**- thyroid carcinoma
- **UCEC**- uterine corpus endometrial carcinoma

## **Introduction**

Cancer is one of the top causes of mortality worldwide; GLOBOCAN estimates that 19.3 million new cancer cases and 10 million deaths will be reported in 2020 (Sung et al., 2021). Several researchers have worked tirelessly over the last few decades to develop novel cures and treatments to fight against the deadly disease (Pucci et al., 2019). Traditional therapies like chemotherapy, radiation, and surgery are the most commonly used treatments (Arruebo et al., 2011). These radiation-based therapies have negative consequences on the patient's health and survival (Altun and Sonkaya, 2018;Pucci et al., 2019;Dilalla et al., 2020). To circumvent the drawbacks of conventional medicines, new treatment regimens have been developed, including targeted cancer therapies, adoptive T cell therapy, immune checkpoint inhibitor-based therapies, immunomodulators, interferons and oncolytic viruses (Padma, 2015;Dine et al., 2017;Esfahani et al., 2020;Franzin et al., 2020;Hemminki et al., 2020). Cancer immunotherapy have resulted in considerable outcomes and increase the life duration of many patients suffering from various solid tumours (Amin et al., 2020;Ruiz-Patino et al., 2020).

The central pillars of immunotherapy are immune checkpoint inhibitors and chimeric antigen receptor (CAR) T cells. These therapies are entirely dependent on T-lymphocytes (T cells), which recognise tumor-associated peptides displayed on the tumor cell surface by human leukocyte antigens (HLA) (Waldman et al., 2020). HLAs are the highly complex and polymorphic genes in the human genome, situated on chromosome 6. Class-I HLA alleles interact with the CD8<sup>+</sup> T cell receptors to activate T cells which further induce several

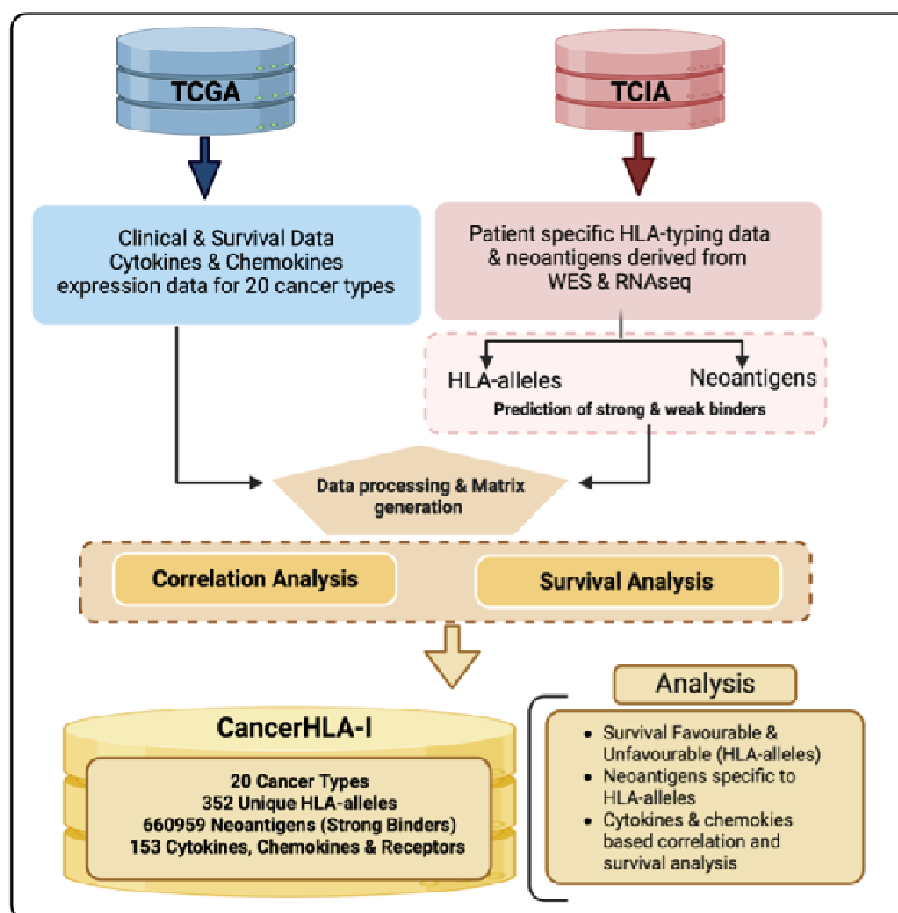
immune responses to knock off the tumor cells from our system (Buhrman and Slansky, 2013;Engels et al., 2013;Chan et al., 2018;He et al., 2019). The immunotherapies are entirely based on the T cells, which identify tumor-associated peptides presented by human leukocytes antigens (HLA) on the infected cell surface. Recently, scientists majorly focused on HLA-dependent therapies, including CD8+ T cell therapy, tumor-infiltrating lymphocytes (TILs) therapy, and TCR-engineered T cells (TCR-Ts) and neoantigen-based therapy (Sun et al., 2021;Yarmarkovich et al., 2021) to treat cancer patients. HLA-dependent treatments are more effective and efficient than traditional chemotherapies. HLA-peptide binding is critical for determining cancer immunogenicity.

In the past number of repositories such as TCIA (Charoentong et al., 2017), TSNadb (Wu et al., 2018), NEPdb (Xia et al., 2021), dbpepNeo (Tan et al., 2020), Ovirusdb (Lathwal et al., 2020), CancerTope (Gupta et al., 2016) have been developed for designing immunotherapy. However, these databases lack with information of impact of class-I HLA-alleles, neobinders and cytokines on the survival of cancer patients. To stratify patient-specific therapy, HLA-typing, neoantigens, and binding affinity must be identified. It is now possible to detect patient-specific HLA-alleles with the help of latest technologies and the availability of sequencing data. To construct patient-specific therapy, genetic information such as HLA-alleles, neoantigens, HLA-peptide binding affinity, and immune response need to be considered. In the piolet study, we have developed a resource which provide patient-specific information from public repositories such as (TCGA and TCIA) and analysed patient survival based on HLA-alleles, as well as the correlation of the amount of neobinders specific to HLA-alleles with overall survival in various cancer types. Furthermore, we used correlational analysis to better understand the role of chemokines, cytokines, and their receptors in cancer patient prognosis. We combine all of the aforementioned analysis of 20 types of cancers onto a single user-friendly resource named as “CancerHLA-I” (<https://webs.iiitd.edu.in/raghava/cancerhla1/>).

## **Material and Methods**

### ***Overall study design***

The complete architecture of the study is depicted in Figure 1



**Figure 1: Complete architecture of the study including dataset collection, analysis and website development**

### *Collection and pre-processing of datasets*

At first, we gathered the genomic and clinical data for 8346 patients with 20 different types of cancer (BLCA, BRCA, CESC, CRC, GBM, HNSC, KICH, KIRC, KIRP, LIHC, LUAD, LUSC, OV, PAAD, PRAD, READ, SKCM, STAD, THCA, and UCEC) from The Cancer Genome Atlas (TCGA) (Tomczak et al., 2015) and The Cancer Immunome Atlas TCIA (Charoentong et al., 2017) repositories. We obtain the control excess dataset patient-specific HLA-typing data and neoantigens data for 20 types of cancers available in TCIA [with the approval of dbGap (Project No. 17674)]. Furthermore, RNA-seq expression data of cytokines, chemokines, and their receptors was obtained using TCGA assembler 2.0. The expression profiles were then transformed into log<sub>2</sub> values after addition of 1.0 as a constant

number to each of expression value. The survival information covers vital status and overall survival time (OS).

### ***Mean-overall and univariate survival analysis***

For each cancer type, we first built a binary matrix based on the presence or absence of HLA-alleles. Each column represents HLA-alleles, and each row represents samples/patients. We used individual's survival data to calculate mean overall survival (MOS) based on the presence or absence of an HLA-allele. After that, we computed the difference in MOS (depending on presence/absence). Cox proportional hazard (Cox-PH) regression models were used in the current study to identify HLA-alleles linked with cancer patient survival. The “survival” package in R (V.3.5.1) was used for univariate analysis. The cox regression coefficient greater than 0 indicates that the presence of an HLA-allele which affects survival (unfavourable), whereas less than 0 indicates that the presence of alleles increases survival (favourable). For each allele, we calculated the Hazard Ratio (HR) and 95% CI (Confidence Interval).  $HR > 1$  denotes high-risk HLA-alleles, while  $HR < 1$  depicts low-risk alleles; however,  $HR = 1$  has no effect on survival. Furthermore, the log-rank test and p-value were conducted to determine the significant distribution of low-risk and high-risk patients. To calculate the predictive performance of models, we used the Concordance index (C). In this study univariate survival analysis performed based on HLA-alleles, cytokines, and chemokines genes for 20 each cancer type.

### ***Correlation analysis***

We extracted the strong binding neoantigens/epitopes corresponding to each HLA-allele for each cancer type using the MHCflurry 2.0 software (O'Donnell et al., 2020). We classify neoepitopes as strong or weak binders using the MHCflurry software's binding affinity (BA) percentile, where neoantigens with  $BA < 2$  are considered strong binders and  $BA > 2$  taken as weak binders. Following that, we built a count matrix with the number of strong binders matching to each HLA-allele and cancer type. We calculate the correlation coefficient between overall survival and the number of strong binders for each HLA-allele in order to understand the impact of number of binders on the survival using Pearson correlation test. The correlation coefficient and p-value ( $< 0.05$ ) demonstrate the significance of the number of HLA-binding neoepitopes on cancer patient survival. Moreover, we conducted correlation

analysis using the expression values of cytokines, chemokines genes for each cancer type. We used survival data and the expression of 153 cytokines, chemokines, and their receptors.

### ***Database implementation***

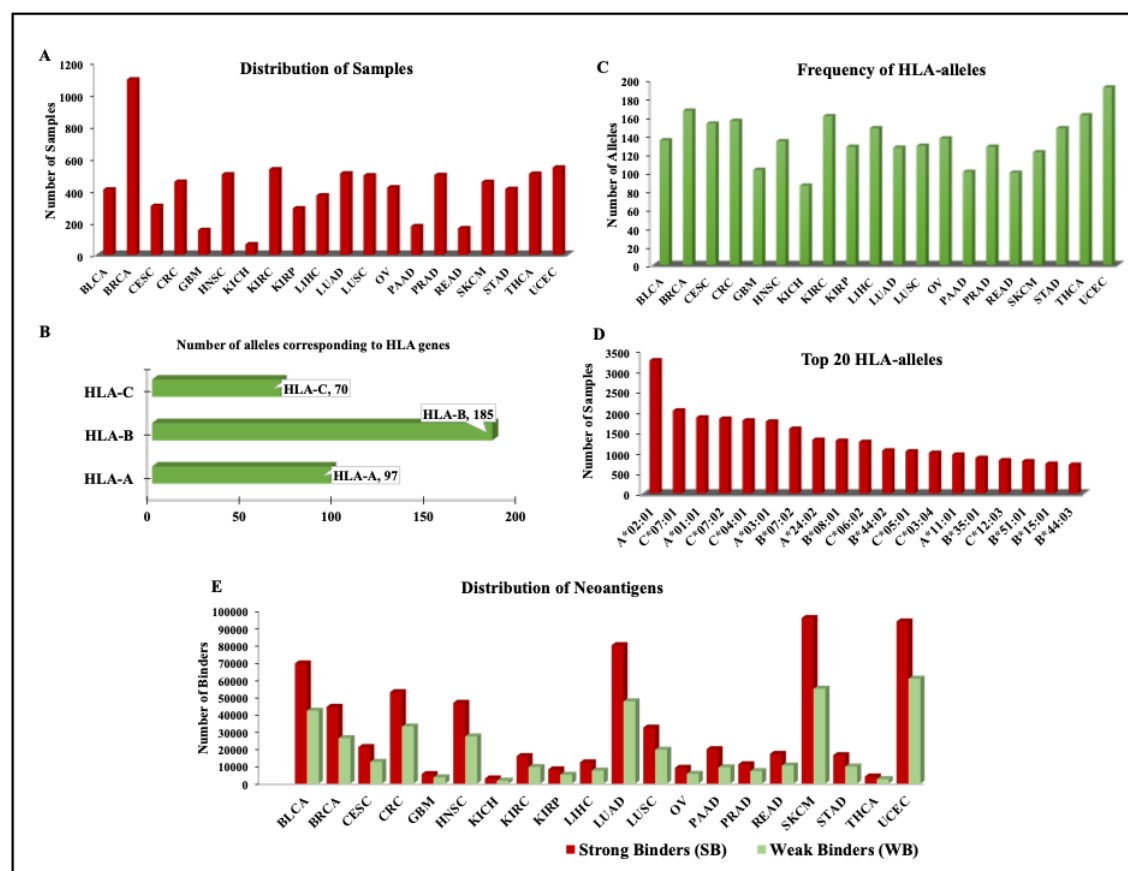
The web interface of CancerHLA-I (<https://webs.iiitd.edu.in/raghava/cancerhla1/>) was developed using MySQL and hosted on linux based apache server. In order to create the interactive user interface, we developed responsive framework made up of HTML, CSS, and JavaScript. To improve the data view, the user interface is responsive, which means that the web interface recognises the user device and modifies its structure and shape in accordance with the device resolution. With the help of this functionality, the interface is adaptable to a wide range of devices and browsers with various screen resolutions. The web-site can be searched using variety of devices (smartphones or tablets) and browsers (Google Chrome, Mozilla Firefox, and Safari).

## **Results**

### ***Statistical analysis of data of CancerHLA-I***

CancerHLA-I incorporates data on cancer associated HLA-alleles, neoantigens, cytokines, chemokine, and their relationship with the overall survival of 20 cancer types. The genomic and clinical information of 8346 cancer patients were downloaded from TCGA and TCIA repositories and processed to build this resource. In Figure 2, we have provided the description of 20 types of cancer, with number of samples, HLA-alleles, total number neoantigens, strong and weak binders. As shown in Figure 2B, HLA-B acquire maximum number of alleles i.e., 185 followed by HLA-A (97 alleles) and HLA-C (70 alleles). We observed that in the case of Uterine Corpus Endometrial Carcinoma (UCEC) highest number (i.e., 192 HLA-alleles) are reported, whereas Kidney Chromophobe (KICH) reported the lowest number i.e., 86 HLA-alleles (See Figure 2C). In addition, we reported top-20 most frequent HLA-alleles whose frequency distribution is maximum among cancer patients (See Figure 2D). Moreover, we have obtained more than 90,000 strong binder i.e., neobinders corresponding to Skin Cutaneous Melanoma (SKCM) and Uterine Corpus Endometrial Carcinoma (UCEC) cancers. However, we have less than 5000 neobinders for Kidney chromophobe (KICH) and Thyroid Carcinoma (THCA) cancer types (See Figure 2E). The complete data used in this analysis is stored in big MySQL table that can be searched and

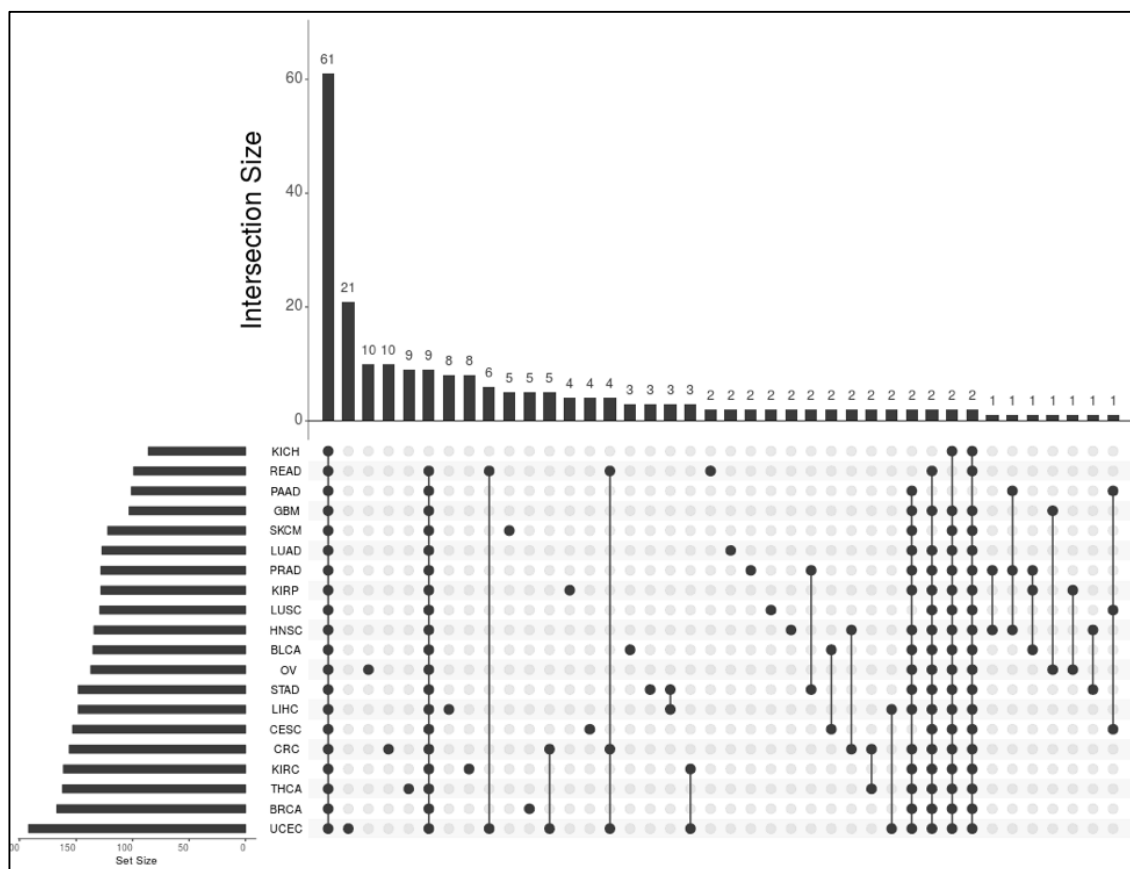
browsed by various categories defined in the database such as cancer type, HLA-allele, neoantigen, HR, p-value, cytokine/chemokine, correlation coefficient etc.



**Figure 2: Complete distribution of dataset in 20 cancer types with number of samples, frequency of alleles, HLA-alleles corresponding to class-I genes, most frequent HLA-alleles, and distribution of neoantigens**

Of note, we have generated the UpSet plot to understand the distribution of HLA-alleles in different cancer types. These overlapping and exclusive HLA-alleles among each cancer type (See Figure 3) are the ones that can serve the basis for explaining the molecular heterogeneity and similarity among the different cancer types. It also provides the potential insights into the progression of disease.





**Figure 3: UpSet plot of top 40 interactions including both common and unique HLA-alleles among each cancer type**

### *HLA-based prognostic biomarkers*

At first, we have combined all the cancer types to performed survival analysis in order to check the overall impact of presence/absence of HLA alleles on the survival of the cancer patients. As shown in the Supplementary Table S1, we have not observed any significant trend by combining all the cancer types due to tumor heterogeneity. So we have generated binary matrix corresponding to each cancer type based on the presence and absence of HLA-alleles. With the utility of survival package we have performed survival analysis computed to identify the high risk and low risk HLA-alleles corresponding to each cancer type. In Table 1, we have reported only those whose HLA-alleles whose presence influences the survival of more than one type of cancer patients. For instance, presence of HLA-A\*02:01 allele reduces the survival of kidney cancer (HR=5.46 with p-value=0.03) and skin cancer patients (HR=1.36 and p-value=0.02). We observed that presence of HLA-A\*02:01, HLA-A\*68:01,

HLA-B\*52:01 and HLA-C\*03:02 is significantly associated with the poor prognosis (with HR>2 and p-value<0.05) in different cancer patients, as shown in Table 1. Moreover, the presence of certain HLA-alleles significantly improves the survival rate of cancer patients for instance, HLA-C\*14:02, HLA-C\*12:03, HLA-A\*03:01 significantly improve the survival rate and act as good prognostic markers in Bladder urothelial carcinoma (BLCA), Stomach Adenocarcinoma (STAD), Liver Hepatocellular Carcinoma (LIHC) and Glioblastoma Multiforme (GBM) (See Table 1).

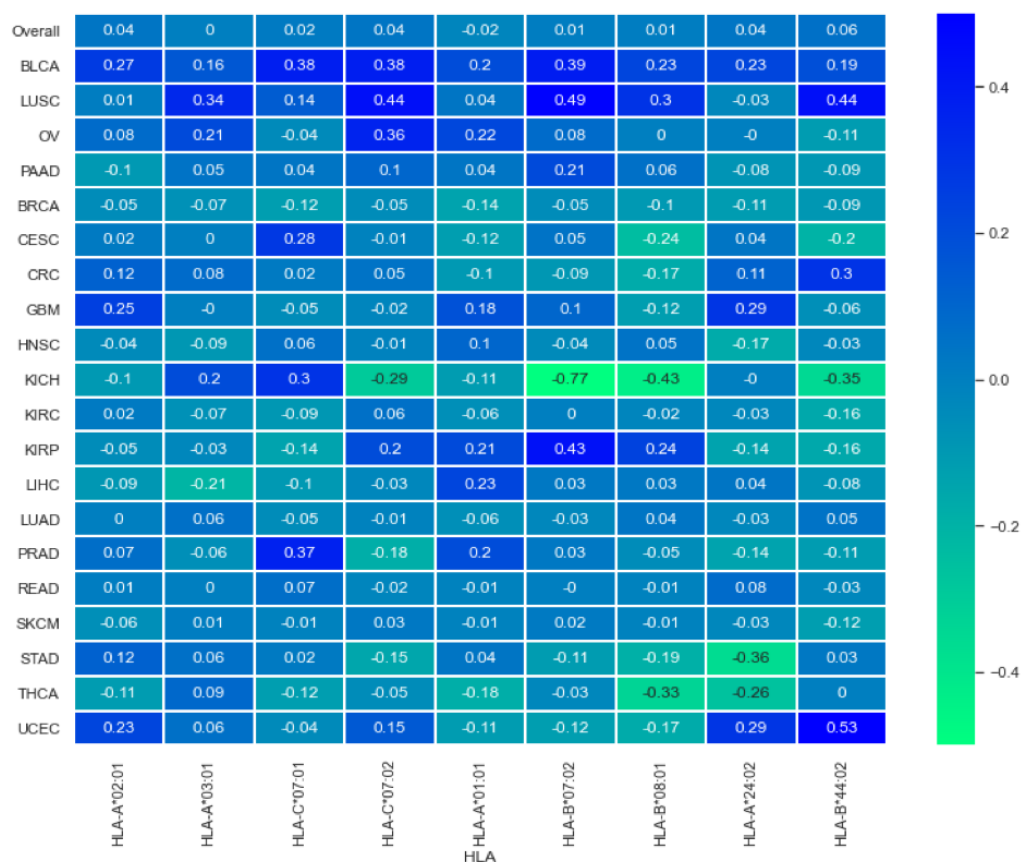
**Table 1: Results of univariate survival analysis based on the presence/absence of HLA-alleles in different type of cancers**

Cancer	HLA-allele	Allele Present (No. of Patients)	Allele Absent (No. of Patients)	Hazard Ratio (95% CI)	p-value	Concordance Index
KICH	HLA-A*02:01	26	39	5.462 (1.135-26.299)	0.034	0.72
SKCM		203	251	1.36 (1.038-1.781)	0.025	0.528
CRC	HLA-A*68:01	24	431	2.466 (1.237-4.918)	0.01	0.532
OV		21	399	1.936 (1.101-3.402)	0.022	0.51
GBM	HLA-B*44:03	16	138	1.805 (1.025-3.181)	0.041	0.528
LIHC		46	324	1.665 (1.041-2.663)	0.033	0.529
OV		31	389	1.583 (1.019-2.459)	0.041	0.512
THCA	HLA-B*52:01	25	480	4.057 (1.155-14.254)	0.029	0.62
SKCM		16	438	2.183 (1.154-4.132)	0.016	0.515
HNSC		18	483	2.113 (1.149-3.886)	0.016	0.514
LIHC	HLA-C*03:02	16	354	2.268 (1.104-4.658)	0.026	0.517
BRCA		96	997	1.659 (1.025-2.685)	0.039	0.52
LIHC		43	327	1.651 (1.034-2.636)	0.036	0.533
LUAD	HLA-C*07:01	140	367	1.366 (0.998-1.869)	0.051	0.538
LUSC		139	356	1.365 (1.018-1.829)	0.037	0.526

<b>LUAD</b>		51	456	1.668 (1.075-2.587)	0.022	0.518
<b>BRCA</b>	HLA-C*12:03	113	980	1.655 (1.042-2.627)	0.033	0.543
<b>GBM</b>		22	132	0.527 (0.29-0.936)	0.029	0.532
<b>KIRP</b>	HLA-A*03:01	74	215	1.884 (1.007-3.526)	0.044	0.531
<b>LIHC</b>		72	298	0.635 (0.397-1.015)	0.046	0.526
<b>BLCA</b>	HLA-C*14:02	15	392	0.140 (0.02-1.001)	0.048	0.517
<b>STAD</b>		20	390	0.315 (0.1-0.988)	0.048	0.516

### ***HLA-I neobinders based correlation analysis***

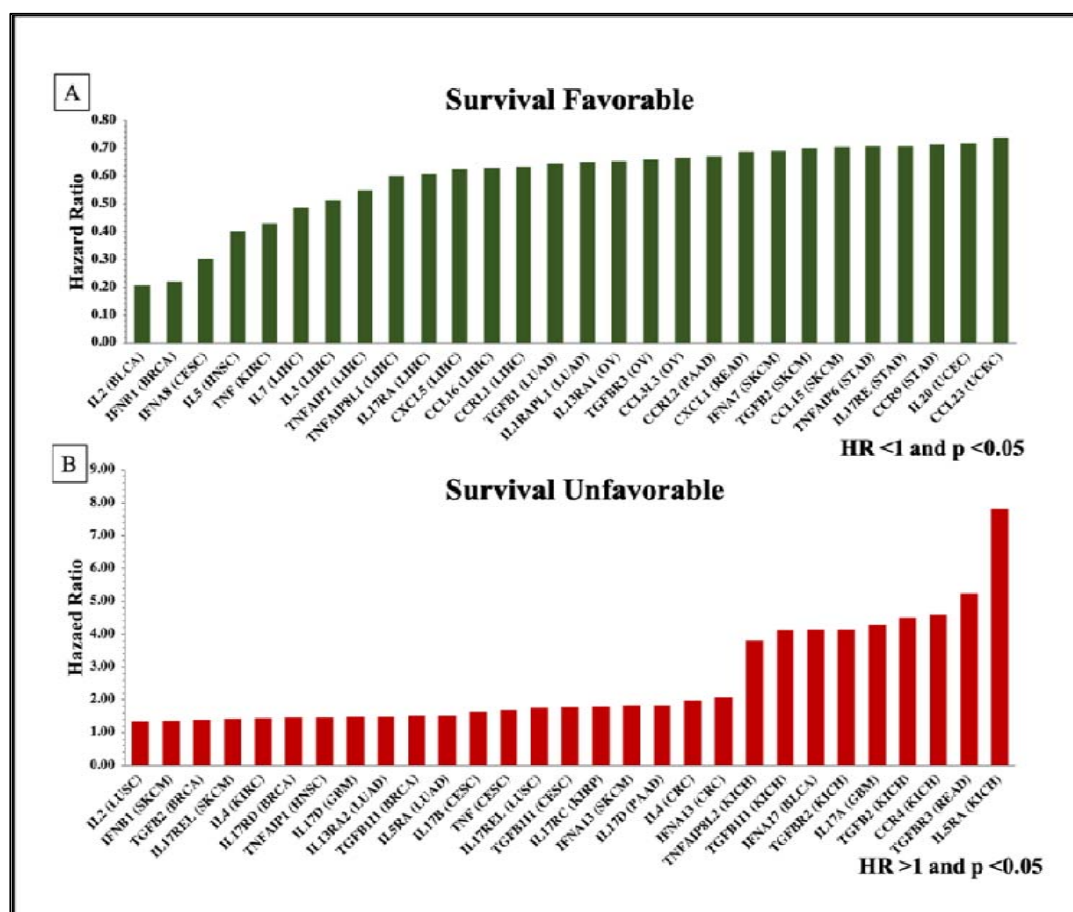
In this study, we have used MHCflurry 2.0 (O'Donnell et al., 2020) software for the prediction of neoepitopes having strong binding potential with the class-I HLA alleles. We have identified strong HLA-specific neobinders for each cancer type. In order to understand the correlation or impact of number of neobinders with the survival of cancer patients, we performed Pearson correlation analysis. Figure 4, shows the correlation values for the nine HLA-alleles (HLA-A\*02:01, HLA-A\*03:01, HLA-C\*07:01, HLA-C\*07:02, HLA-A\*01:01, HLA-B\*07:02, HLA-B\*08:01, HLA-A\*24:02, and HLA-B\*44:02) present in most of the samples and all cancer types. At first, we have computed the correlation between the neobinders irrespective of cancer type by combining all the data files. We observed that, the overall impact on combining binders for all the cancer types is very less or negligible (See Figure 4). While, some of the neobinders corresponding to particular HLA-allele have very high positive as well as negative correlation with the specific cancer types. For instance, HLA-B\*07:02 neobinders have very high negative correlation ( $r = -0.77$ ) with the survival of KICH patients, while on the opposite side it shows positive correlation of ( $r = 0.49$ ) with the LUSC patients. In the case of BLCA, LUSC and OV most of the nine alleles shows positive correlation with the overall survival. However, in the case of KICH some of the alleles shows positive association and some of them shows negative association with the survival. This means binders corresponding to HLA-alleles have favourable and unfavourable impact depends upon the cancer types.



**Figure 4: Correlation analysis based on the number of HLA-I neobinders with the overall survival of cancer patients**

### *Cytokines based prognostic biomarkers*

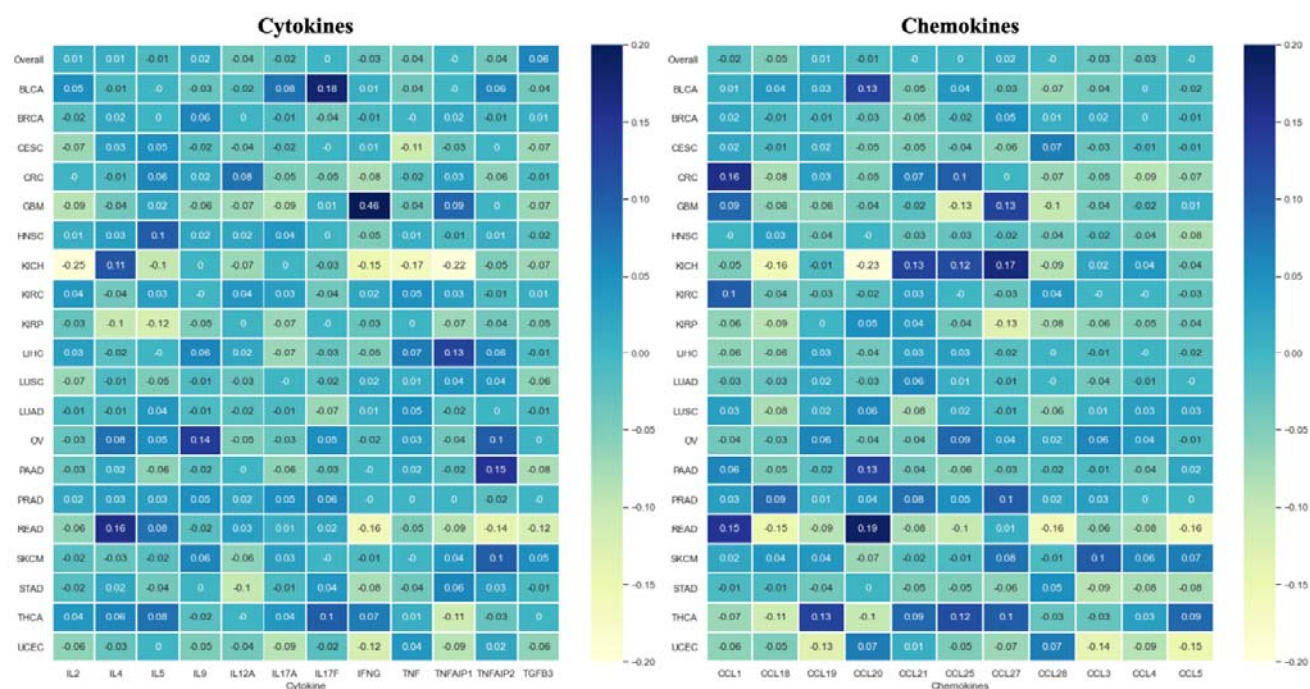
In order to understand the prognostic role of cytokines and chemokines we have performed univariate survival analysis using their expression profiles. In the Figure 5, we have reported those cytokine and chemokines whose expression significantly impact the survival rate of cancer patients. We observed that high expression of *IL2*, *IFNB1*, *IFNA8*, *IL5* cytokines are having very good impact on the survival of different cancer patients ( $HR < 0.4$  and  $p\text{-value} < 0.05$ ). Whereas, *IL5RA*, *TGFBR3*, *CCR4*, *TGFB2*, *IL17A* are highly associated with the poor survival rate in KICH, READ, and GBM patients ( $HR > 4$  and  $p\text{-value} < 0.05$ ). The complete analysis of all other cytokines and chemokines is available on CancerHLA-I server.



**Figure 5: Univariate survival analysis shows survival favourable (HR<1) and unfavourable (HR>1) cytokines, chemokines and their receptors in different cancer types**

Moreover, we have done correlation analysis by considering the gene expression of cytokine, chemokines and their receptors. The heatmap shows (Figure 6) the correlation of overall survival with the expression of some of the cytokines and chemokines in 20 cancer types. The darker blue colour shows the positive correlation, whereas light yellow colour depicts the negative correlation. We observed that cytokine IFNG have very high and significant positive correlation with the survival rate of GBM patients, higher expression of IL9 cytokine is associated with positive correlation in BLCA and OV cancer patients. Whereas, cytokine *IL2*, *TNFAIP1*, and *TNF* are associated with the negative correlations with the survival of KICH patients. In case of chemokines, we observed that *CCL1* (CRC, KIRC and READ), *CCL20* (BLCA, READ, and PAAD), *CCL27* (GBM, KICH and THCA) have positive correlation

with the overall survival. However, the over expression of *CCL18*, *CCL28*, *CCLA*, *CCL5* associated with negative correlation with most of the cancer types (See Figure 6).



**Figure 6: Correlation analysis based on the expression profile of cytokines & chemokines with the overall survival of cancer patients**

### Utility of CancerHLA-I

CancerHLA-I can be interactively browsed and searched in a variety of different ways to satisfy the query of the user. The homepage of CancerHLA-I website provides a simple search page, where users can search query in the database for specific cancer type, HLA-allele, neoantigens, cytokine/chemokine, and its survival association (See Figure 7). The advanced search page in the database provides customized search facility for user defined query using Boolean expressions (AND/OR). We compiled the data in a tabular form corresponding to each cancer type for easy and efficient access. The advantage of the browsing facility is that users can quickly obtain all the results by clicking onto specific entry under concerned category. Moreover, 'Help' page on the website provides detailed visualization of the usage of the CancerHLA-I database. The data can be downloaded as a tab-delimited, comma separated, JSON, PDF and XLSX file formats.

**(A) User can select search options from the home page of website**

**(B) Results are displayed in responsive tables**

**(C) Details of CancerHLA-ID**

**(D) Displayed neoepitopes corresponding to CancerHLA-ID**

**Figure 7: Schematic illustration of the utility of CancerHLA-I (A). Simple search page of the database CancerHLA-I, where users can select fields for querying against the database. Here we have shown the results for the HLA-A\*01:01 in the quick search box (B). Complete results against this HLA-alleles are displayed in a responsive table (C). Users can select ID to get detailed information against this query (D). Users can also click on the neoepitopes and download HLA specific binders in the different file formats**

## Discussion and Conclusion

Class-I (HLA-A, HLA-B and HLA-C) molecules are essential for immunosurveillance and cancer immunotherapy (Sabbatino et al., 2020; Hazini et al., 2021). It is crucial to present tumor specific peptides or neoantigens via HLA-alleles for the detection and killing of tumor cells by our immune system (van den Bulk et al., 2018; Jiang et al., 2019; Peng et al., 2019; Zhang et al., 2021). However, the loss of the functions of class-I HLA molecules exhibit escape mechanism by different cancer types (Garrido et al., 2016; Dhatchinamoorthy et al., 2021; Hazini et al., 2021). Studies also report that upregulation of class-I non-classical

HLA molecules play important role in the cancer immune escape (Bukur et al., 2012; Kochan et al., 2013). Due to mutations at genetic and epigenetic levels loss of heterozygosity occurs in HLA genes at chromosome 6 in non-small cell lung carcinoma and colorectal cancer patients (Hazini et al., 2021; Zhang and Sjoblom, 2021). A recent study reveal that the presence of specific HLA-allele can alter the effect of therapy in different cancers. Researchers observed the presence of HLA-A\*03 allele in kidney cancer patients results into reduced survival and poor respond against immune checkpoint blockade (ICI) therapy (Naranbhai et al., 2022). In addition, previous studies report that HLA-B\*55 and HLA-A\*01 significantly improves the survival while HLA-B\*50 allele reduces the survival rate in melanoma patients (Dhall et al., 2020). Moreover, mutations in type-I and II interferon pathway genes also effect the survival of cancer patients. Interleukins such as *IL-6*, *IL-11*, *IL-1*, and *TGF- $\beta$*  induces cancer cell proliferation and progression (Esquivel-Velazquez et al., 2015). Studies reveals that cytokines play important role in the regulation of tumor microenvironment.

Therefore, it is crucial to understand the prognostic role of HLA-alleles and cytokines in order to know the impact of efficacy of cancer immunotherapy. In the present work, we have conducted a study to investigate the connections of Class-I HLA alleles with cancer patient survival in order to aid researchers. We performed pan-cancer analysis on more than 8000 patients in 20 different cancer types. The dataset used in this study obtained from TCGA and TCIA repositories. We used survival data to determine the association between the presence or absence of the 352 unique HLA-alleles. Furthermore, we investigate the relationship between cytokine expression with the overall survival in cancer patients. We observed that the HLA-A\*02:01, HLA-A\*68:01, HLA-B\*52:01, HLA-C\*03:02 associated with the poor survival and HLA-C\*14:01, HLA-C\*12:03 and HLA-A\*03:01 improves the survival of cancer patients. Moreover, correlation analysis revealed the positive and negative association of neobinders with the survival rate of the cancer patients. Moreover, we also identified some of the high expression levels of cytokines *IL2*, *IFNG*, *IFNB*, *TNF* significantly improves the survival of cancer patients while some of the cytokines like *IL5*, *IL17*, *CCR4*, *TGF* reduces the survival rate significantly. We incorporated overall results in the highly interactive web-based platform for the analysis and identification of cancer-specific biomarkers. We have integrated user-friendly browsing, searching, and analysis modules in our resource “CancerHLA-I” (<https://webs.iiitd.edu.in/raghava/cancerhla/>) for easy data retrieval, data



comparison, and examination. We anticipate that our research yields promising novel HLA and cytokines based biomarkers for improved cancer immunotherapy and treatment.

## **Funding Source**

The current work has received grant from the Department of Bio-Technology (DBT), Govt. of India, India.

## **Conflict of interest**

The authors declare no competing financial and non-financial interests.

## **Authors' contributions**

AD and GPSR collected and processed the datasets. AD, HK, SP and GPSR conceived the idea and analysed the results. AD and SP created the back-end of the web server the front-end user interface. AD, SP and GPSR penned the manuscript. GPSR conceived and coordinated the project. All authors have read and approved the final manuscript.

## **Acknowledgements**

Authors are thankful to the Department of Bio-Technology (DBT) and Department of Science and Technology (DST-INSPIRE) for fellowships and the financial support and Department of Computational Biology, IIITD New Delhi for infrastructure and facilities.

## **References**

- Altun, I., and Sonkaya, A. (2018). The Most Common Side Effects Experienced by Patients Were Receiving First Cycle of Chemotherapy. *Iran J Public Health* 47, 1218-1219.
- Amin, S., Baine, M.J., Meza, J.L., and Lin, C. (2020). Association of Immunotherapy With Survival Among Patients With Brain Metastases Whose Cancer Was Managed With Definitive Surgery of the Primary Tumor. *JAMA Netw Open* 3, e2015444.
- Arruebo, M., Vilaboa, N., Saez-Gutierrez, B., Lambea, J., Tres, A., Valladares, M., and Gonzalez-Fernandez, A. (2011). Assessment of the evolution of cancer treatment therapies. *Cancers (Basel)* 3, 3279-3330.
- Buhrman, J.D., and Slansky, J.E. (2013). Improving T cell responses to modified peptides in tumor vaccines. *Immunol Res* 55, 34-47.
- Bukur, J., Jasinski, S., and Seliger, B. (2012). The role of classical and non-classical HLA class I antigens in human tumors. *Semin Cancer Biol* 22, 350-358.

- Chan, K.F., Gully, B.S., Gras, S., Beringer, D.X., Kjer-Nielsen, L., Cebon, J., McCluskey, J., Chen, W., and Rossjohn, J. (2018). Divergent T-cell receptor recognition modes of a HLA-I restricted extended tumour-associated peptide. *Nat Commun* 9, 1026.
- Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., Hackl, H., and Trajanoski, Z. (2017). Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep* 18, 248-262.
- Dhall, A., Patiyal, S., Kaur, H., Bhalla, S., Arora, C., and Raghava, G.P.S. (2020). Computing Skin Cutaneous Melanoma Outcome From the HLA-Alleles and Clinical Characteristics. *Front Genet* 11, 221.
- Dhatchinamoorthy, K., Colbert, J.D., and Rock, K.L. (2021). Cancer Immune Evasion Through Loss of MHC Class I Antigen Presentation. *Front Immunol* 12, 636568.
- Dilalla, V., Chaput, G., Williams, T., and Sultanem, K. (2020). Radiotherapy side effects: integrating a survivorship clinical lens to better serve patients. *Curr Oncol* 27, 107-112.
- Dine, J., Gordon, R., Shames, Y., Kasler, M.K., and Barton-Burke, M. (2017). Immune Checkpoint Inhibitors: An Innovation in Immunotherapy for the Treatment and Management of Patients with Cancer. *Asia Pac J Oncol Nurs* 4, 127-135.
- Engels, B., Engelhard, V.H., Sidney, J., Sette, A., Binder, D.C., Liu, R.B., Kranz, D.M., Meredith, S.C., Rowley, D.A., and Schreiber, H. (2013). Relapse or eradication of cancer is predicted by peptide-major histocompatibility complex affinity. *Cancer Cell* 23, 516-526.
- Esfahani, K., Roudaia, L., Buhlaiga, N., Del Rincon, S.V., Papneja, N., and Miller, W.H., Jr. (2020). A review of cancer immunotherapy: from the past, to the present, to the future. *Curr Oncol* 27, S87-S97.
- Esquivel-Velazquez, M., Ostoa-Saloma, P., Palacios-Arreola, M.I., Nava-Castro, K.E., Castro, J.I., and Morales-Montor, J. (2015). The role of cytokines in breast cancer development and progression. *J Interferon Cytokine Res* 35, 1-16.
- Franzin, R., Netti, G.S., Spadaccino, F., Porta, C., Gesualdo, L., Stallone, G., Castellano, G., and Ranieri, E. (2020). The Use of Immune Checkpoint Inhibitors in Oncology and the Occurrence of AKI: Where Do We Stand? *Front Immunol* 11, 574271.
- Garrido, F., Aptsiauri, N., Doorduijn, E.M., Garcia Lora, A.M., and Van Hall, T. (2016). The urgent need to recover MHC class I in cancers for effective immunotherapy. *Curr Opin Immunol* 39, 44-51.
- Gupta, S., Chaudhary, K., Dhanda, S.K., Kumar, R., Kumar, S., Sehgal, M., Nagpal, G., and Raghava, G.P. (2016). A Platform for Designing Genome-Based Personalized Immunotherapy or Vaccine against Cancer. *PLoS One* 11, e0166372.
- Hazini, A., Fisher, K., and Seymour, L. (2021). Deregulation of HLA-I in cancer and its central importance for immunotherapy. *J Immunother Cancer* 9.
- He, Q., Jiang, X., Zhou, X., and Weng, J. (2019). Targeting cancers through TCR-peptide/MHC interactions. *J Hematol Oncol* 12, 139.
- Hemminki, O., Dos Santos, J.M., and Hemminki, A. (2020). Oncolytic viruses for cancer immunotherapy. *J Hematol Oncol* 13, 84.
- Jiang, T., Shi, T., Zhang, H., Hu, J., Song, Y., Wei, J., Ren, S., and Zhou, C. (2019). Tumor neoantigens: from basic research to clinical applications. *J Hematol Oncol* 12, 93.
- Kochan, G., Escors, D., Breckpot, K., and Guerrero-Setas, D. (2013). Role of non-classical MHC class I molecules in cancer immunosuppression. *Oncoimmunology* 2, e26491.
- Lathwal, A., Kumar, R., and Raghava, G.P.S. (2020). OvirusTdb: A database of oncolytic viruses for the advancement of therapeutics in cancer. *Virology* 548, 109-116.

- Naranbhai, V., Viard, M., Dean, M., Groha, S., Braun, D.A., Labaki, C., Shukla, S.A., Yuki, Y., Shah, P., Chin, K., Wind-Rotolo, M., Mu, X.J., Robbins, P.B., Gusev, A., Choueiri, T.K., Gulley, J.L., and Carrington, M. (2022). HLA-A\*03 and response to immune checkpoint blockade in cancer: an epidemiological biomarker study. *Lancet Oncol* 23, 172-184.
- O'donnell, T.J., Rubinsteyn, A., and Laserson, U. (2020). MHCflurry 2.0: Improved Pan-Allele Prediction of MHC Class I-Presented Peptides by Incorporating Antigen Processing. *Cell Syst* 11, 42-48 e47.
- Padma, V.V. (2015). An overview of targeted cancer therapy. *Biomedicine (Taipei)* 5, 19.
- Peng, M., Mo, Y., Wang, Y., Wu, P., Zhang, Y., Xiong, F., Guo, C., Wu, X., Li, Y., Li, X., Li, G., Xiong, W., and Zeng, Z. (2019). Neoantigen vaccine: an emerging tumor immunotherapy. *Mol Cancer* 18, 128.
- Pucci, C., Martinelli, C., and Ciofani, G. (2019). Innovative approaches for cancer treatment: current perspectives and new challenges. *Ecancermedicalscience* 13, 961.
- Ruiz-Patino, A., Arrieta, O., Cardona, A.F., Martin, C., Raez, L.E., Zatarain-Barron, Z.L., Barron, F., Ricaurte, L., Bravo-Garzon, M.A., Mas, L., Corrales, L., Rojas, L., Lupinacci, L., Perazzo, F., Bas, C., Carranza, O., Puparelli, C., Rizzo, M., Ruiz, R., Rolfo, C., Archila, P., Rodriguez, J., Sotelo, C., Vargas, C., Carranza, H., Otero, J., Pino, L.E., Ortiz, C., Laguado, P., Rosell, R., and Clicap (2020). Immunotherapy at any line of treatment improves survival in patients with advanced metastatic non-small cell lung cancer (NSCLC) compared with chemotherapy (Quijote-CLICaP). *Thorac Cancer* 11, 353-361.
- Sabbatino, F., Liguori, L., Polcaro, G., Salvato, I., Caramori, G., Salzano, F.A., Casolaro, V., Stellato, C., Col, J.D., and Pepe, S. (2020). Role of Human Leukocyte Antigen System as A Predictive Biomarker for Checkpoint-Based Immunotherapy in Cancer Patients. *Int J Mol Sci* 21.
- Sun, Y., Li, F., Sonnemann, H., Jackson, K.R., Talukder, A.H., Katailiha, A.S., and Lizee, G. (2021). Evolution of CD8(+) T Cell Receptor (TCR) Engineered Therapies for the Treatment of Cancer. *Cells* 10.
- Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 71, 209-249.
- Tan, X., Li, D., Huang, P., Jian, X., Wan, H., Wang, G., Li, Y., Ouyang, J., Lin, Y., and Xie, L. (2020). dbPepNeo: a manually curated database for human tumor neoantigen peptides. *Database (Oxford)* 2020.
- Tomczak, K., Czerwinska, P., and Wiznerowicz, M. (2015). The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)* 19, A68-77.
- Van Den Bulk, J., Verdegaal, E.M., and De Miranda, N.F. (2018). Cancer immunotherapy: broadening the scope of targetable tumours. *Open Biol* 8.
- Waldman, A.D., Fritz, J.M., and Lenardo, M.J. (2020). A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nat Rev Immunol* 20, 651-668.
- Wu, J., Zhao, W., Zhou, B., Su, Z., Gu, X., Zhou, Z., and Chen, S. (2018). TSNadb: A Database for Tumor-specific Neoantigens from Immunogenomics Data Analysis. *Genomics Proteomics Bioinformatics* 16, 276-282.
- Xia, J., Bai, P., Fan, W., Li, Q., Li, Y., Wang, D., Yin, L., and Zhou, Y. (2021). NEPdb: A Database of T-Cell Experimentally-Validated Neoantigens and Pan-Cancer Predicted Neopeptides for Cancer Immunotherapy. *Front Immunol* 12, 644637.
- Yarmarkovich, M., Marshall, Q.F., Warrington, J.M., Premaratne, R., Farrel, A., Groff, D., Li, W., Di Marco, M., Runbeck, E., Truong, H., Toor, J.S., Tripathi, S., Nguyen, S., Shen, H., Noel, T., Church, N.L., Weiner, A., Kendsersky, N., Martinez, D.,

- Weisberg, R., Christie, M., Eisenlohr, L., Bosse, K.R., Dimitrov, D.S., Stevanovic, S., Sgourakis, N.G., Kiefel, B.R., and Maris, J.M. (2021). Cross-HLA targeting of intracellular oncoproteins with peptide-centric CARs. *Nature* 599, 477-484.
- Zhang, X., and Sjoblom, T. (2021). Targeting Loss of Heterozygosity: A Novel Paradigm for Cancer Therapy. *Pharmaceuticals (Basel)* 14.
- Zhang, Z., Lu, M., Qin, Y., Gao, W., Tao, L., Su, W., and Zhong, J. (2021). Neoantigen: A New Breakthrough in Tumor Immunotherapy. *Front Immunol* 12, 672356.